

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: 2 0 0 3 年 4 月 2 日

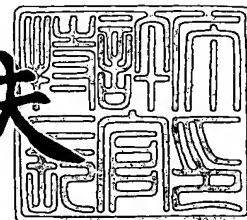
出 願 番 号
Application Number: 特 願 2 0 0 3 - 0 9 9 4 6 4
[ST. 10/C]: [J P 2 0 0 3 - 0 9 9 4 6 4]

出 願 人
Applicant(s): キヤノン株式会社

2 0 0 4 年 4 月 1 9 日

特許庁長官
Commissioner,
Japan Patent Office

今 井 康 夫



出証番号 出証特 2 0 0 4 - 3 0 3 2 8 6 2

【書類名】 特許願

【整理番号】 251776

【提出日】 平成15年 4月 2日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/00

【発明の名称】 D N A プローブ設計装置及びD N A プローブ設計のための
情報処理方法

【請求項の数】 14

【発明者】

【住所又は居所】 東京都大田区下丸子 3 丁目 3 0 番 2 号 キヤノン株式会
社内

【氏名】 吉井 裕人

【特許出願人】

【識別番号】 000001007

【氏名又は名称】 キヤノン株式会社

【代理人】

【識別番号】 100076428

【弁理士】

【氏名又は名称】 大塚 康德

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100112508

【弁理士】

【氏名又は名称】 高柳 司郎

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100115071

【弁理士】

【氏名又は名称】 大塚 康弘

【電話番号】 03-5276-3241

【選任した代理人】

【識別番号】 100116894

【弁理士】

【氏名又は名称】 木村 秀二

【電話番号】 03-5276-3241

【手数料の表示】

【予納台帳番号】 003458

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 0102485

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 DNAプローブ設計装置及びDNAプローブ設計のための情報処理方法

【特許請求の範囲】

【請求項1】 ターゲットの塩基配列データを含む第1の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第1カウント工程と、

前記第1の塩基配列データ群とは区別されるべき第2の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第2カウント工程と、

前記第1及び第2カウント工程によって保持された頻度情報を、前記複数の部分塩基配列を基準に対比可能に表示する表示工程と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを特定し、特定された部分塩基配列に基づいてプローブ候補を形成する形成工程とを備えることを特徴とするDNAプローブ設計のための情報処理方法。

【請求項2】 前記表示工程は前記第1及び第2カウント工程によって保持された頻度情報を、前記ターゲットの塩基配列を横軸とするグラフで表示し、

前記指示操作は、前記グラフ上の所望の位置を指示することによりなされることを特徴とする請求項1に記載のDNAプローブ設計のための情報処理方法。

【請求項3】 ターゲットの塩基配列データを含む第1の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第1カウント工程と、

前記第1の塩基配列群とは区別されるべき第2の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第2カウント工程と、

前記第1カウント工程による頻度が第1所定値を超え、前記第2カウント工程による頻度が第2所定値を下回る部分塩基配列を探索する探索工程と、

前記探索工程で探索された部分塩基配列に基づいてプローブ候補を形成する形成工程とを備えることを特徴とするDNAプローブ設計のための情報処理方法。

【請求項 4】 前記複数の部分塩基配列は、前記ターゲットの塩基配列データより所定数の塩基を抽出してなる塩基配列を、その先頭位置を順次ずらしながら取得することにより得られたものであることを特徴とする請求項 1 乃至 3 のいずれかに記載の DNA プローブ設計のための情報処理方法。

【請求項 5】 前記第 1 グループの塩基配列データ群はターゲットとする生物種の複数の多型を含む塩基配列データであり、前記第 2 グループの塩基配列データは前記ターゲットとする生物種以外の生物種の複数の多型を含む塩基配列データであることを特徴とする請求項 1 乃至 4 のいずれかに記載の DNA プローブ設計のための情報処理方法。

【請求項 6】 前記形成工程において形成されたプローブ候補についてプローブの融解温度を算出し、算出された融解温度に基づいてプローブセットに用いるべきプローブ候補を選択する第 1 選択工程を更に備えることを特徴とする請求項 1 乃至 5 のいずれかに記載のプローブ設計のための情報処理方法。

【請求項 7】 前記形成工程において形成されたプローブ候補について 2 次構造作成可能性を算出し、その算出結果に基づいてプローブセットに用いるべきプローブ候補を選択する第 2 選択工程を更に備えることを特徴とする請求項 1 乃至 6 のいずれかに記載のプローブ設計のための情報処理方法。

【請求項 8】 前記形成工程において形成されたプローブ候補についてマッチング度を計算し、マッチング度に基づいてプローブセットに用いるべきプローブ候補を選択する第 3 選択工程を更に備えることを特徴とする請求項 1 乃至 7 のいずれかに記載のプローブ設計のための情報処理方法。

【請求項 9】 ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント手段と、

前記第 1 の塩基配列データ群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント手段と、

前記第 1 及び第 2 カウント手段によって保持された頻度情報を、前記複数の部分塩基配列を基準に対比可能に表示する表示手段と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを特定し、特定された部分塩基配列に基づいてプローブ候補を形成する形成手段とを備えることを特徴とするDNAプローブ設計装置。

【請求項10】 ターゲットの塩基配列データを含む第1の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第1カウント手段と、

前記第1の塩基配列群とは区別されるべき第2の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第2カウント手段と、

前記第1カウント手段による頻度が第1所定値を超え、前記第2カウント手段による頻度が第2所定値を下まわる部分塩基配列を探索する探索手段と、

前記探索手段で探索された部分塩基配列に基づいてプローブ候補を形成する形成手段とを備えることを特徴とするDNAプローブ設計装置。

【請求項11】 請求項1乃至8のいずれかに記載のプローブ設計のための情報処理方法をコンピュータに実行させるための制御プログラム。

【請求項12】 請求項1乃至8のいずれかに記載のプローブ設計のための情報処理方法をコンピュータに実行させるための制御プログラムを格納した記憶媒体。

【請求項13】 請求項1乃至8のいずれかに記載のプローブ設計方法で設計された核酸プローブを含むDNAマイクロアレイ。

【請求項14】 請求項13に記載のDNAマイクロアレイを用いた核酸検査方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明はいわゆるDNAマイクロアレイ等を用いた核酸配列解析システムに好適なオリゴヌクレオチドプローブの設計技術に関するものである。

【0002】

【従来の技術】

従来より、特許文献 1 または特許文献 2 にあるように、DNA マイクロアレイを用いて遺伝子の発現やシーケンスを決定するシステムが知られている。これらの特許文献に開示されたシステムは、cDNA をスポットして作成される DNA マイクロアレイとは異なり、試料とハイブリダイズさせるためのプローブをあらかじめ設計する必要がある。うまく適切なプローブを設計できれば、非常に高い確率で試料の中に存在する塩基配列の断片に対する情報が得られる。

【0 0 0 3】

このシステムにおいて、プローブとして用いられる塩基配列の長さは、長くても 1 0 0 ベース以上になることは珍しく、短いものでは数個の塩基配列からなるものまである。すなわち、上記特許文献で開示されたシステムでは、cDNA に比べて非常に短い塩基配列のプローブを用いて特定の塩基配列を捕らえることになる。このため、DNA 中のプローブとして用いる部分の塩基配列の特異性が非常に高い必要がある。

【0 0 0 4】

上記のような特異性の高い部分を選び出すための従来の選択方法では、一般的な配列に対する特異性評価が行われる。例えば、人間のゲノム由来の DNA を対象とした DNA マイクロアレイを作成しようとした場合、人間の全ゲノム塩基配列に対する特異性を調べて、特異性の高い部分をプローブ塩基配列として選んでいた。

【0 0 0 5】

【特許文献 1】

特開平 1 0 - 2 7 2 0 0 0 号公報

【特許文献 2】

特開平 1 1 - 1 8 7 9 0 0 号公報

【0 0 0 6】

【発明が解決しようとする課題】

しかしながら、従来の選択方法では、非常に似ている塩基配列が検体の中に入っていて、かつ、それら類似した塩基配列があるグループに属する塩基配列と別のグループに属する塩基配列を含む場合、各塩基配列がそのグループに属するか

を判定することは非常に困難であるという問題が存在していた。より具体的には、感染症などの判定を行う際に、同種の菌の別種類の株の DNA 塩基配列には同じようなハイブリ強度を示し、かつ、別の種の菌の DNA 塩基配列には異なるハイブリ強度を示すようなプローブを見つけ出すことが非常に困難であるという問題があった。

【 0 0 0 7 】

また、従来の方法では、非常に似ている塩基配列の中で、生物種に特異的な場所、かつ、多型には共通な場所を探そうとする場合、マルチプルアラインメントなどの方法を用いて、対象としている生物種の全ての多型の塩基配列を揃えて表示し、人間が目で見えて妥当な部分を選んでいた。そして、この従来法では、人間のミスが混入したり、人によるバラツキが生じていた。

【 0 0 0 8 】

本発明は、上記課題に鑑みてなされたものであり、正確かつ、再現性のあるプローブ設計を可能とすることを目的とする。

【 0 0 0 9 】

【課題を解決するための手段】

上記の問題点を解決するために本発明による DNA プローブ設計のための情報処理方法は、

ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント工程と、

前記第 1 の塩基配列データ群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント工程と、

前記第 1 及び第 2 カウント工程によって保持された頻度情報を、前記複数の部分塩基配列を基準に対比可能に表示する表示工程と、

ユーザによる指示操作に応じて前記複数の部分塩基配列の少なくとも一つを特定し、特定された部分塩基配列に基づいてプローブ候補を形成する形成工程とを備える。

【0 0 1 0】

また、本発明の他の態様による DNA プローブ設計のための情報処理方法は、ターゲットの塩基配列データを含む第 1 の塩基配列データ群について、該ターゲットの塩基配列データから得られる複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 1 カウント工程と、

前記第 1 の塩基配列群とは区別されるべき第 2 の塩基配列データ群について、前記複数の部分塩基配列の各々の出現数をカウントして得られる頻度情報を保持する第 2 カウント工程と、

前記第 1 カウント工程による頻度が第 1 所定値を超え、前記第 2 カウント工程による頻度が第 2 所定値を下まわる部分塩基配列を探索する探索工程と、

前記探索工程で探索された部分塩基配列に基づいてプローブ候補を形成する形成工程とを備える。

【0 0 1 1】**【発明の実施の形態】**

以下、添付の図面に基づいて本発明の好適な実施形態を説明する。

【0 0 1 2】**〔プローブ設計方法の説明〕**

図 2 は本実施形態によるプローブ設計方法が適用される情報処理装置の構成を示すブロック図である。本実施形態のプローブ設計方法は、外部記憶装置 2 0 1、中央処理装置（CPU）2 0 2、メモリ 2 0 3、入出力装置 2 0 4 から構成される装置に実装される。すなわち、一般的なパーソナルコンピュータ、ワークステーション等の実装可能である。

【0 0 1 3】

図 2 において、外部記憶装置 2 0 1 は、本実施形態のプローブ設計方法を実現するプログラムや、各種塩基配列データ及びパラメータ（DNA（オリゴヌクレオチド）プローブ長や融解温度等）を保持する。また、本実施形態によって選択されたプローブ配列そのものを保持する機能を持つ。中央処理装置（CPU）2 0 2 はプローブ設計のプログラムを実行したり、すべての装置の制御を行ったりする。メモリ 2 0 3 は中央処理装置（CPU）2 0 2 が使用するプログラム、

及びサブルーチンやデータを一時的に記憶する。入出力装置 204 は、ディスプレイ、キーボード、ポインティングデバイス等を含み、ユーザとのインタラクションを行う。多くの場合、本実施形態のプローブ設計方法を実現するプログラム実行のトリガはこの入出力装置を介してユーザが出す。また、ユーザが結果を見たり、プログラムのパラメータ制御をこの入出力装置を介して行う。

【0014】

図1は、実施形態によるプローブ設計方法の処理手順を説明するフローチャートである。101は自己塩基配列データであり、例えばターゲット塩基配列107がある特定の菌の特定株の塩基配列であった場合、ターゲット塩基配列107と同種の菌の種々の株が有する塩基配列のデータからなる。102は自己頻度テーブル作成ステップであり、自己塩基配列データ101に含まれる部分塩基配列の頻度をカウントして自己頻度テーブルを作成する。103はターゲット塩基配列と競合する競合塩基配列データであり、自己塩基配列データ101の菌種とは異なる種々の菌種の種々の株が有する塩基配列のデータからなる。104は競合頻度テーブル作成ステップであり、競合塩基配列データ103に含まれる部分塩基配列の頻度をカウントして競合頻度テーブルを作成するステップである。

【0015】

108はプローブ評価ステップであり、自己頻度テーブル105と競合頻度テーブル106を用いて、ターゲット塩基配列107からプローブ候補を導き出す。109はプローブセット選択ステップであり、プローブ評価ステップ109の結果得られたプローブ候補より最適なプローブセットを選択する。110は最終的に得られた最適プローブを示す。

【0016】

図1に示すプローブ設計処理において、入力ターゲット塩基配列107であり、出力は最適プローブ110である。

【0017】

図1の各データ及び処理について詳述する前に、本発明のバックグラウンドとなる知識を説明する。図3はDNAマイクロアレイ上のハイブリダイゼーションの様子を示した図である。生体内でほとんどの場合、DNA塩基列は2重らせん

構造をしていて、その2本鎖の間の結合は塩基間の水素結合で実現されている。一方、RNA塩基列は1本で存在する場合が多い。塩基の種類はDNAの場合はACGTの4種類、RNAの場合はACGUの4種類であり、それぞれ水素結合ができる塩基対はA-T (U)、G-Cのペアとなっている。ハイブリダイゼーションとは、1本鎖状態の核酸分子同士がある部分の塩基配列を介して部分的に結合する状態をいい、本実施形態で想定している反応は、図3の上側の基板にくっついた核酸分子（プローブ）の方が下側の資料中にある核酸分子より短い。よって、資料中に存在する核酸分子がプローブ塩基配列を含む場合は、このハイブリダイゼーション反応はうまくいき、試料中のターゲット核酸分子はトラップされることとなる。

【0018】

次に、図4を用いて感染症の菌を特定するDNAマイクロアレイの原理を説明する。図4で示したDNAマイクロアレイは、例えば、黄色ブドウ球菌を特定する目的で作られていると仮定する。図4の左側は、黄色ブドウ球菌野生株由来の処理系列であり、右側は大腸菌野生株由来の処理系列である。例えば、左側は黄色ブドウ球菌に感染した患者の血液を処理する流れで、右側は大腸菌に感染した患者の血液を処理する流れだと考えてよい。

【0019】

どちらも基本的には同じ処理を行う。つまり、まず初めに例えば菌感染患者の血液や、痰などからDNAを抽出する（401、411）。この際に、一般的には、患者の体細胞由来の人間のDNAも含まれる可能性がある。抽出されたDNAが少ない場合、PCRなどの方法で増幅を行う。この際に蛍光物質もしくは蛍光物質を結合させることができる物質を標識として混入させるのが一般的である（402、412）。

【0020】

増幅をしない場合は、抽出されたDNAを用いて、相補鎖を作りながら蛍光物質もしくは蛍光物質を結合させることができる物質を標識として混入させる、または、そのまま直接抽出されたDNAに蛍光物質もしくは蛍光物質を結合させることができる物質を標識として付加させる（403、413）。

【0021】

通常、PCR増幅を行う場合、感染症の菌特定が目的であれば、いわゆる16sといわれるリボゾームRNA（16s rRNA）を構成する塩基配列の部分を増幅するのが一般的である。この場合、図4の左側の黄色ブドウ球菌のPCRプライマーと右側の大腸菌のPCRプライマーはほとんど同じものを使うこととなる。より具体的には、どんな菌の16s rRNAをコーディングしている部分でも増幅させることができるプライマーセットを用いて、マルチプレックスPCRを行うことが好ましい。この場合、結果的に、図4における右と左のどちらのハイブリ溶液（404、414）も複数の種類の塩基配列を含むものとなる。この理由は次の図で詳しく述べる。

【0022】

これに対して、より詳しい塩基配列解析を行いたい場合は、例えば、黄色ブドウ球菌用のPCRプライマーセット、大腸菌用のPCRプライマーセットを別々に設定する。この場合、仮に菌のゲノムの特定部分だけを選択的に増幅するようにプライマーを設定すれば、ハイブリ溶液に含まれる塩基配列の種類は非常に限定される。それでも通常、自然界に存在する菌の株は数種類に及ぶので、ハイブリ溶液に存在する塩基配列の種類が1種類となることは稀である。

【0023】

ここで黄色ブドウ球菌を判定する目的のために設計されたDNAマイクロアレイが正しく動作するならば、ハイブリ溶液404ではスポットがポジティブに反応し（405）、右側のハイブリ溶液414では、スポットがネガティブに反応する（415）。

【0024】

これと全く同じように、大腸菌の存在を判定する目的のために設計されたDNAマイクロアレイが正しく動作するならば、ハイブリ溶液404ではスポットがネガティブに反応し、ハイブリ溶液414ではスポットがポジティブに反応する。もちろん、いろんな菌に対してそれぞれ特異的に反応する数種類のスポットを同時に並べたDNAマイクロアレイを用いて、感染菌の判定を行ってもかわまない。

【0 0 2 5】

次に、図 5 を用いて、図 4 のハイブリ溶液に複数の種類の塩基配列が存在する理由を説明する。通常、自然界に存在する菌は、突然変異を頻繁に起こす。その結果、淘汰を経て生き残った主要な数種類の株が同時に存在することがある。例えば、院内感染などで問題を起こす菌株は、通常は薬剤耐性がないはずの菌が、突然変異を起こすことによって薬剤耐性を獲得することで出現する。薬剤耐性を獲得した結果、殺菌努力を行っている衛生的な環境でも旺盛な繁殖力を持つ菌が出現したりする。このように、自然界に存在する同一の菌の塩基配列は、数種類のバリエーションをもつと考えるのが妥当である。

【0 0 2 6】

図 5 に示したのは、黄色ブドウ球菌に Mu50 と MW2 の 2 種類の株のゲノム構造である。それぞれの株のゲノムの総塩基数はそれぞれ 2,878,040 と 2,820,462 と異なっている。なお、図 5 では、左から右への方向を 5' 末端から 3' 末端方向とし、これを順方向として塩基配列が示されている。更に 16s のリボゾーム RNA (16s rRNA) をコーディングしている部位は、Mu50 の場合は順方向 2 箇所、逆方向 3 箇所の合計 5 箇所であるのに対して、MW2 では順方向 3 箇所、逆方向 3 箇所の合計 6 箇所存在する。これらの 16s rRNA の各部位の塩基配列はそれぞれ非常に似ているが、全く同じではない。つまり、調べようとしている菌感染患者の体の中に存在する菌の株の種類がたとえ 1 種類であったとしても、図 4 のような一般的な処理でハイブリ溶液を調製すると複数種類の塩基配列がハイブリ溶液の中に存在することになる。この複数の塩基配列に対して同じようなハイブリ強度を示すプローブを設計することが本実施形態のプローブ設計方法の目的である。

【0 0 2 7】

この目的を達成するために本実施形態のプローブ設計方法は、図 1 に示した通り、ターゲット塩基配列と同じグループに属する塩基配列のセット（自己塩基配列データ 1 0 1）と、そのグループに競合するグループに属する塩基配列のセット（競合塩基配列データ 1 0 3）から別々に頻度テーブルを作成する。図 4 の例で言うと、例えば黄色ブドウ球菌の色々な株の色々な部位の 16s rRNA の塩基配列の集合が自己塩基配列データ 1 0 1 となり、黄色ブドウ球菌以外の例えば大腸菌

や、インフルエンザ菌などの種々の菌の色々な株の色々な部位の16s rRNAの塩基配列の集合が競合塩基配列データ103となる。

【0028】

これらの塩基配列データから頻度テーブルを作成する様子を図6に示す。頻度テーブルの作成においては、塩基配列データに含まれる長さ n （図6では $n=9$ ）の部分配列の出現回数をカウントする。長さ n の塩基配列のバリエーションは4の n 乗なので、図6において、行の数は 4^n となる。なお、図6において出現頻度が低いほど、その部分配列の特異性は高くなるので、例えば頻度にマイナス1をかけた数の特異性とする。

【0029】

すなわち、自己頻度テーブル作成ステップ102は検出対象の菌種の種々の株の16s rRNAの塩基配列を記憶する自己塩基配列データ101を参照して、長さ n の全ての部分塩基配列について出現回数をカウントし、その結果を図6の如くテーブル化し、自己頻度テーブル105とする。同様に、競合頻度テーブル作成ステップ112は検出対象の菌種以外の菌種（検出対象から区別されるべき菌種）の種々の株の16s rRNAの塩基配列を記憶する競合塩基配列データ101を参照して、長さ n の全ての部分塩基配列について出現回数をカウントし、その結果を図6の如くテーブル化し、競合頻度テーブル115とする。

【0030】

次に、図1のターゲット塩基配列107に対する部分列（ n 個の塩基からなる塩基配列）の頻度または特異性を上記頻度テーブルを用いて求める。この様子を示したのが図7である。このターゲット塩基配列107は、例えば、図5におけるMu50株の黄色ブドウ球菌の一番左の16s rRNAコーディング部位（配列X）などとなる。プローブ評価ステップ108では、自己頻度テーブル105及び競合頻度テーブル106を参照して、ターゲット塩基配列107から順次得られる部分列の頻度（特異性）を取得し、評価する。この様子をグラフに示したのが図8である。

【0031】

図8の横軸はターゲット塩基配列の部分配列の位置であり、例えば16s rRNAの

部分を調べたいとすれば、その長さは約 1 5 0 0 程度となる。縦軸は、その位置における部分配列の特異性であり、例えば頻度テーブルの値にマイナス 1 をかけた数である。図 8 において、上段に示したグラフが図 1 における競合頻度テーブル 1 0 6 を元に導いた特異性を表し、下段に示したグラフが図 1 における自己頻度テーブル 1 0 5 を元に導いた特異性を表す。

【 0 0 3 2 】

例えば、図 8 のグラフの場合、配列の先頭から約 2 / 3 の部分、つまり約 1 0 0 0 ベースぐらいの位置で、このターゲット塩基配列の特異性の高い部分が存在する。例えば、このターゲット塩基配列が図 5 の Mu50 株の最初の 16s 部分（配列 X）だったとすると、この配列 X の 1000 ベースあたりの位置で、この配列に特有の塩基配列が存在し、それは、Mu50 株のその他の 16s の部位、または、黄色ブドウ球菌の Mu50 株以外の株の 16s rRNA 部位とは異なる配列であることを意味する。よって、この部分からプローブ候補を選択することは、好ましくない。また、上段のグラフにおいて特異性の低い部分は、他の種の菌との識別ができないことを意味しており、その部分からプローブ候補を選択することも好ましくない。

【 0 0 3 3 】

したがって、上段のグラフで山となっており（異種菌間で特異性が高い）、かつ下段のグラフで山となっていない部分（同種菌の異種株間で特異性が低い）よりプローブ候補を選択する。このことで、判定すべき菌とは異なる菌種に対しては特異性が高く（＝頻度が低く）、判定すべき菌の色々な変異体に対しては特異性が低い（頻度が高い）プローブを選ぶことができる。すなわち、図 8 のように、各特異性を複数の部分塩基配列を基準に対比可能に表示することにより、ユーザは容易に適切なプローブ候補を選択できる。

【 0 0 3 4 】

この結果、同じ菌であれば、どのような株のどのような部位の 16s rRNA コーディング配列がハイブリ溶液に含まれていても強いハイブリ反応を示し、かつ、他の菌のどのような株のどのような部位の 16s rRNA コーディング配列がハイブリ溶液に含まれていても弱いハイブリ反応を示すプローブが選ばれることとなる。

【 0 0 3 5 】

なお、本発明のプローブ設計方法は、感染症の判別にのみ適用できるものではなく、一般的に同一と判断すべき塩基配列の対象の中に若干のバリエーションが存在するものなら何でも適用できる。例えば人間の個体判別によく用いられるMHCなどの判定などにも適用できる。

【0036】

次に、図1のプローブセット選択ステップ109について説明する。図1のプローブ評価ステップ108で高い評価を得たプローブの中で、長さを揃えてプローブセットとするのが一番単純なプローブセット選択の方法である。しかしながら、ハイブリダイゼーション反応はプローブ塩基配列の長さよりも、その融解温度で決まる場合が一般的である。よって、例えば標準プローブ長 n （本例では $n=24$ とする）を設定しておいて、その長さの前後数塩基の許容範囲（本例では ± 2 とする）より得られる各プローブについて融解温度を求め、出来るだけ融解温度を一定にするように採用すべきプローブを決定することによって、より高い品質のプローブセットが得られる。

【0037】

なお、塩基配列の融解温度は、その配列の塩基の混入率で求める方法や、いわゆるNearest Neighbor法と呼ばれる連続する2つの塩基配列の並び方から求める方法などが知られており、これらの方法を用いて融解温度を算出することができる。

【0038】

また、塩基配列の長さが20を超えてくると、2次構造を形成して、プローブとして適切でない場合が発生する。これを防ぐために、例えばMichael Zuker氏の考案した"Calculating nucleic acid secondary structure." (Current Opinion in Structural Biology, 10, 303-310, (2000)) に記載の方法などを用いて2次構造形成の可能性を計算し、2次構造を形成しやすいプローブは排除するようにしてもよい。

【0039】

また、図1のプローブ評価ステップで説明した頻度テーブルを用いる方法は、あくまで特異性の目安としてプローブ候補を選ぶ方法であるので、実際に選ばれ

たプローブ群が似かよったものとなる可能性がある。したがって、プローブ候補がどの程度お互いに似ているかを、お互いのマッチングを取ることで調べ、似かよったプローブ候補を排除することが好ましい。例えば、N個のプローブ候補からプローブセットを選択するとすると、 $N(N-1)/2$ 個のマッチングを取り、プローブ同士の塩基配列がお互いにどれくらい異なっているかを評価して、異なる塩基数がより多いプローブセットを選ぶようにする。このことによって、より高い品質のプローブセットを選択することができる。なお、この方法は、いわゆるクロスハイブリダイゼーションを防ぐ方法と呼ばれる。

【0040】

[プローブ設計装置のより詳細な説明]

図9～図11を用いて本実施形態のプローブ設計プログラムの動きを説明する。プローブ設計プログラムの動作はまずターゲット生物グループの選定(901)から始まる。例えば、感染症の原因菌特定のためのプローブを設計しようとしたとすると、ターゲット生物グループが、菌やウイルス、真菌などのゲノム情報が塩基配列データベース906から選ばれる。図9において、塩基配列データベース906は、NCBIなどの公共のデータベースや、自社のイントラネット上に構築された塩基配列データベースなどで、現在入手可能な最大量のデータが蓄積されたものを意味する。一方、ターゲット塩基配列データベース907は、ターゲット生物グループの選定(901)で選ばれた生物種のゲノム情報しか含まない。例えば、本プログラムを人間の体質判定のためのプローブ設計に適用した場合は、ターゲット塩基配列データベース907に格納される塩基配列は、例えばMHCのDRB1の部分の全アレルの情報などとなる。

【0041】

次に、ターゲット生物種の選定を行う(902)。ターゲット生物種を選定すると、ターゲット塩基配列データベース907に含まれる塩基配列が自己塩基配列データと競合塩基配列データに分けられる。すなわち、ターゲット塩基配列データベース907に含まれる情報から、ターゲット生物種の複数の多型、複数のゲノム部位に対応する自己塩基配列データ908が抽出されるとともに、ターゲット塩基配列データベース907に含まれる、ターゲット生物種以外の生物種の

複数の多型、複数のゲノム部位に対応する競合塩基配列データ 909 が抽出される。

【0042】

次に、選ばれたターゲット生物種を元に自己頻度テーブル 910 と競合頻度テーブル 911 を作成する (903)。この頻度テーブルの作成 (903) のとき、通常は、ターゲットとするゲノム領域も設定する。例えば菌種の判別を行うプローブを設計する場合には、16s rRNA の部分選ばれたりする。

【0043】

より具体的に述べると、図 4 で述べたとおり、通常、DNA マイクロアレイを用いた実験を行う場合、PCR でターゲットとなる核酸を増幅する。この時、PCR プライマーで挟んだ領域のみが増幅されるので、ターゲット塩基配列データベース 907 の中の PCR で増幅される部分のみを用いて頻度テーブルを作成する。本プログラムを人間の体質判定のためのプローブ設計に適用した場合は、例えばターゲットとする DNA 領域も MHC の DRB1 の部分などに予め設定されるので、この DNA 領域を設定する必要はない。また、例えば MHC の DRB1 の解析であれば、たんに質発現に差を示さない 3 桁までのアレルを同一種として扱う。なお、ターゲットとする核酸領域、または PCR 増幅領域の指定は、通常、プログラムのプロパティとして設定し、UI (ユーザインターフェース) を使ってその都度設定することはない。

【0044】

次にターゲット塩基配列の選定を行い (904)、そのターゲット塩基配列の部分列群の特異性を自己頻度テーブル 910 と競合頻度テーブル 911 を用いて評価し、プローブを選定する (905)。

【0045】

図 10 は、ターゲット生物種の選定 (902) からターゲット塩基配列の選定 (904) までを行うユーザインターフェースの例を示したものである。まず、ターゲット菌種欄 1001 にターゲット菌種としての菌の一覧が表示される。ここで、例えば黄色ブドウ球菌 (*Staphylococcus aureus*) を選ぶと、株表示欄 1002 に黄色ブドウ球菌の色々な株の色々な部位の 16s rRNA を示す一覧が表示さ

れる。ここまでのインタラクションで、図9の902、903（ターゲット生物種の選定と頻度テーブルの作成）が実行される。株表示欄1002に示されている多型としての複数の株は、図5に示した株である。例えば、図10のように、株の名前、ゲノム上の位置、方向を示すことによって、それぞれの塩基配列の識別ができるようにする。

【0046】

株表示欄1002の一覧の中で1つの配列を選ぶと、ターゲット塩基配列の選定（904）が実行される。図10に示したように黄色ブドウ球菌Mu50株の先頭の16s rRNAを選んだとすると、ターゲット塩基配列として図5の配列Xが選ばれたことになる。そして、デザインボタン1003を押すとデザイン画面が表示される。

【0047】

図11は実際のデザイン画面の例を示した図である。1101から1104はグラフを示し、横軸は、例えば図10のようなUIで選ばれたターゲット塩基配列の位置を意味し、それぞれのグラフはそのターゲット塩基配列のそれぞれの位置の部分塩基配列に対する値をグラフ化したものである。グラフ1102とグラフ1103は、それぞれ図8の上段下段のグラフに対応しており、グラフ1102はそれぞれの位置の部分配列の競合塩基配列データに対する特異性をグラフ化したもの、グラフ1103はそれぞれの位置の部分配列の自己塩基配列データに対する特異性をグラフ化したものとなる。また、グラフ1101はそれぞれの位置の部分配列の人間のゲノムに対する特異性を示す。グラフ1104はそれぞれの位置の部分配列を先頭とする所定個数の塩基配列（本例では24個の塩基からなる塩基配列（24ベースの塩基配列））の融解温度を示す。

【0048】

図8で説明したように、グラフ1102で山になっていて、グラフ1103で谷になっている場所でプローブを設定するのがよい。1105は情報欄であり、現在のターゲット生物種や各種のパラメータを表示する。なお、本実施形態ではデフォルトの塩基長として24が設定されており、これを基準にグラフ1104の融解温度が計算されている。1106は、設計したプローブのリストで、その

位置は、ライン 1107 (点線) で表示される。1108 の実線は “カレント” 位置を示し、現在注目している位置で、その位置に対応する部分塩基配列 (24 ベースの塩基配列) は 1109 に表示され、その直前の部分の塩基配列は 1110、その直後の部分の塩基配列は 1111 に表示される。本実施形態では前後 10 個の塩基からなる配列が表示される。

【0049】

また、図 11 の UI では、1112 に示すように、ターゲット塩基配列の中で塩基配列を検索する機能を持っている。

【0050】

また、例えば、感染症の原因菌を特定するプローブを設計する場合は、人間のゲノムも含まれてくるので、例えばグラフ 1101 に示すように、人間のゲノムに対するターゲット塩基配列の部分配列の特異性を表示してもよい。また、プローブを設定するために実施例 1 で説明したように融解温度 (T_m) を選択したプローブ間で出来るだけ似通った温度にする方がいい。これを実現するために、例えばグラフ 1104 では、その位置の部分塩基配列の T_m をグラフとして表示する。

【0051】

また、本実施形態では、グラフ 1101 ~ 1104 を参照してユーザが実線を移動することにより、プローブ候補の評価を行なうが、グラフ 1103 で特異性が低く、グラフ 1102 で特異性が高い部分を自動的に抽出してユーザに提示するようにしてもよい。例えば、グラフ 1103 において特異性がある第 1 の閾値よりも低く、グラフ 1102 において特異性がある第 2 の閾値よりも高い部分を抽出してユーザに提示する。なお、上記で設計された塩基配列の相補的な配列も同様にプローブとして用いることができるので、これらを併せて表示する、或いは設計結果として提示するようにしてもよい。

【0052】

[プローブセットの設計例及び実験例]

次に図 12 を用いて、上記実施形態のプローブ設計方法を用いて設計されたプローブを用いた DNA マイクロアレイの実験手順について説明する。

【0053】

1201の“サンプル”とは対象としている核酸が含まれているはずの液体や個体である。例えば感染症の原因菌の特定をする場合、ヒト、家畜等の動物由来の血液、喀痰、胃液、膣分泌物、口腔内粘液等の体液、尿及び糞便のような排出物等細菌が存在すると思われるあらゆる物がサンプルとなる。また、食中毒、汚染の対象となる食品、飲料水及び温泉水のような環境中の水等、細菌による汚染が引き起こされる可能性のある媒体がサンプルとして用いられることもある。さらに、輸出入時における検疫等の動植物も検体としてその対象となる。

【0054】

次に、“生化学的増幅”方法を用いてサンプル1201を増幅する(1202)。例えば感染症の原因菌の特定をする場合、16s rRNA検出用に設計されたPCR反应用プライマーを用いてPCR法によって対象核酸を増幅したり、或いはPCR増幅物を元にさらにPCR反応等を行なって調製したりする。また、PCR以外のLAMP法などの増幅方法により調製してもよい。

【0055】

その後で、生化学的増幅1202により増幅されたサンプル、またはサンプル1201そのものに、可視化のために各種標識法により標識する(ラベル混入1203)。この標識物質としては、通常Cy3, Cy5, Rodaminなどの蛍光物質が用いられる。また、生化学的増幅1202において標識分子を混入することもある。

【0056】

こうして標識分子が付加された核酸は、DNAマイクロアレイ1204とハイブリダイゼーション反応(1205)を行う。この様子は、図3に示した通りである。例えば感染症の原因菌の特定をする場合、DNAマイクロアレイ1204は菌に特異的なプローブを基板に固定したものとなる。ここで、各菌に対応したプローブの設計は、例えば16s rRNAをコーディングしているゲノム部分より、上記で説明した方法で行う。なお、DNAマイクロアレイ1204のプローブを固定する担体(基板)は、ガラス基板、プラスチック基板、シリコンウェハー等の平面基板が考えられる。また、凹凸のある三次元構造体、ビーズのような球状の

もの、棒状、紐状、糸状のもの等を用いても、本発明の実現形態、効果には影響ない。

【0057】

通常、基板の表面はプローブDNAの固定化が可能のように処理したものが使用される。特に、表面に化学反応が可能となるように官能基を導入した物は、ハイブリダイゼーション反応の過程でプローブが安定に結合している為に、再現性の点で好ましい形態である。本実施形態で用いられる固定化方法は、例えば、マレイミド基とチオール（-SH）基との組み合わせを用いる例が挙げられる。即ち核酸プローブの末端にチオール（-SH）基を結合させておき、固相表面がマレイミド基を有するように処理しておくことで、固相表面に供給された核酸プローブのチオール基と固相表面のマレイミド基とが反応して核酸プローブを固定化する。マレイミド基の導入方法としては、まず、ガラス基板にアミノシランカップリング剤を反応させ、次にそのアミノ基とEMCS試薬（N-(6-Maleimidocaproyloxy)succinimide : Dojin社製）との反応によりマレイミド基を導入する。DNAへのSH基の導入は、DNA自動合成機でDNAを合成する際に5'-Thiol-ModifierC6（Glen Research社製）を用いることにより行なうことができる。固定化に利用する官能基の組み合わせとしては、上記したチオール基とマレイミド基の組み合わせ以外にも、例えばエポキシ基（固相上）とアミノ基（核酸プローブ末端）の組み合わせ等が挙げられる。また、各種シランカップリング剤による表面処理も有効であり、該シランカップリング剤により導入された官能基と反応可能な官能基を導入したオリゴヌクレオチドが用いられる。さらに、官能基を有する樹脂をコーティングする方法も利用可能である。

【0058】

ハイブリダイゼーション反応1205を行った後、DNAマイクロアレイ1204の表面を洗浄し、プローブと結合していない核酸を剥がした後で、（通常は）乾燥し、1205の蛍光量を測定する。ここでは、DNAマイクロアレイ1204の基板に励起光を照射し、蛍光強度を測定した画像を得る（1206、1207）。

【0059】

以下、図12を用いて説明した実験の流れを感染症の原因菌特定の目的を想定した具体的実験操作として説明する。なお、本発明にかかわる生物種類判定方法は、以下に述べる感染症の原因菌特定に限ったものではなく、MHCなどの人間の体質判定や、癌などの疾病に関わるDNA、RNAの解析に用いてもよい。

【0060】

<1. プローブDNAの準備>

*Enterobacter cloacae*株検出用Probeとして配列番号59～65に示す核酸配列（I-n）（nは数字）を設計した。具体的には、16s rRNAをコーディングしているゲノム部分より、実施例1、実施例2に述べた方法を用いて設計した。

【0061】

配列番号59～65（相補鎖配列は配列番号137～143）に示したプローブは、DNAマイクロアレイに固定するための官能基として、合成後、定法に従って核酸の5'末端にチオール基を導入した。官能基の導入後、精製し、凍結乾燥した。凍結乾燥したプローブは、-30℃の冷凍庫に保存した。

【0062】

黄色ブドウ球菌、表皮ブドウ球菌、大腸菌、肺炎桿菌、緑膿菌、セラチア菌、肺炎連鎖球菌、インフルエンザ菌、及びエンテロコッカス・フェカリス菌についても同様な手法により以下に示すプローブセットを設計した。

【0063】

黄色ブドウ球菌：配列番号1～9（相補鎖配列は配列番号79～87）、
表皮ブドウ球菌：配列番号10～16（相補鎖配列は配列番号88～94）、
大腸菌：配列番号17～23（相補鎖配列は配列番号95～101）、
肺炎桿菌：配列番号24～29（相補鎖配列は配列番号102～107）、
緑膿菌：配列番号30～37（相補鎖配列は配列番号108～115）、
セラチア菌：配列番号38～43（相補鎖配列は配列番号116～121）、
肺炎連鎖球菌：配列番号44～50（相補鎖配列は配列番号122～128）

インフルエンザ菌：配列番号51～58（相補鎖配列は配列番号129～136）

エンテロコッカス・フェカリス菌：配列番号66～72（相補鎖配列は配列番号144～150）。

【0064】

<2. 検体増幅用PCR Primer の準備>

起炎菌検出用の為の16s rRNA核酸（標的核酸）増幅用PCR Primerとして表2に示す核酸配列を設計した。

【0065】

具体的には、16s rRNAをコーディングしているゲノム部分を特異的に増幅するプローブセット、つまり約1500塩基長の16s rRNAコーディング領域の両端部分で、特異的な融解温度をできるだけ揃えたプライマーを設計した。なお、変異株や、ゲノム上に複数存在する16s rRNAコーディング領域も同時に増幅できるように複数種類のプライマーを設計した。

【0066】

	Primer No.	配列
Forward Primer	F-1	5' GCGGCGTGCCTAATACATGCAAG 3'
	F-2	5' GCGGCAGGCCTAACACATGCAAG 3'
	F-3	5' GCGGCAGGCTTAACACATGCAAG 3'
Reverse Primer	R-1	5' ATCCAGCCGCACCTTCCGATAC 3'
	R-2	5' ATCCAACCGCAGGTTCCTAC 3'
	R-3	5' ATCCAGCCGCAGGTTCCTAC 3'

【0067】

表中に示したPrimerは、合成後、高速液体クロマトグラフィー（HPLC）により精製し、Forward Primerを3種、Reverse Primerを3種混合し、それぞれのPrimer濃度が、最終濃度10 pmol/ μ l となるようにTE緩衝液に溶解した。

【0068】

<3. Enterobacter_cloacae Genome DNA（モデル検体）の抽出>

<3-1. 微生物の培養 & Genome DNA 抽出の前処理>

まず、Enterobacter cloacae 標準株を、定法に従って培養した。この微生物培養液を1.5ml容量のマイクロチューブに1.0ml（OD₆₀₀=0.7）採取し、遠心分離

で菌体を回収した（8500rpm、5min、4℃）。上精を捨てた後、Enzyme Buffer（50mM Tris-HCl：p.H. 8.0、25mM EDTA）300 μ lを加え、ミキサーを用いて再懸濁した。再懸濁した菌液は、再度、遠心分離で菌体を回収した（8500rpm、5min、4℃）。上精を捨てた後、回収された菌体に、以下の酵素溶液を加え、ミキサーを用いて再懸濁した。

Lysozyme	50 μ l	(20 mg/ml in Enzyme Buffer)
N-Acetylmuramidase SG	50 μ l	(0.2 mg/ml in Enzyme Buffer)。

【0 0 6 9】

次に、酵素溶液を加え再懸濁した菌液を、37℃のインキュベーター内で30分間静置し、細胞壁の溶解処理を行った。

【0 0 7 0】

〈3-2. Genome抽出〉

以下に示す微生物のGenome DNA抽出は、核酸精製キット（MagExtractor -Genome-：TOYOBO社製）を用いて行った。

【0 0 7 1】

具体的には、まず、前処理した微生物懸濁液に溶解・吸着液750 μ lと磁性ビーズ40 μ lを加え、チューブミキサーを用いて、10分間激しく攪拌した（ステップ1）。

次に、分離用スタンド（Magical Trapper）にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた（ステップ2）。

次に、洗浄液 900 μ l を加え、ミキサーで5sec程度攪拌して再懸濁を行った（ステップ3）。

次に、分離用スタンド（Magical Trapper）にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた（ステップ4）。

ステップ3、4を繰り返して2度目の洗浄（ステップ5）を行った後、70%エタノール 900 μ l を加え、ミキサーで5sec程度攪拌して再懸濁した（ステップ6）。

【0 0 7 2】

次に、分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブの壁面に集め、スタンドにセットした状態のまま、上精を捨てた (ステップ7)。

ステップ6、7を繰り返して70%エタノールによる2度目の洗浄 (ステップ8) を行った後、回収された磁性粒子に純水 100 μ l を加え、チューブミキサーで10分間攪拌を行った。

次に分離用スタンド (Magical Trapper) にマイクロチューブをセットし、30秒間静置して磁性粒子をチューブ壁面に集め、スタンドにセットした状態のまま、上精を新しいチューブに回収した。

【0 0 7 3】

＜3-3. 回収したGenome DNAの検査＞

回収された微生物 (*Enterobacter cloacae* 株) のGenome DNAは、定法に従って、アガロース電気泳動と260/280nmの吸光度測定を行い、その品質 (低分子核酸の混入量、分解の程度) と回収量を検定した。

【0 0 7 4】

本実施例では、約10 μ gのGenome DNA が回収され、Genome DNAのデグラデーションやrRNAの混入は認められなかった。回収したGenome DNAは、最終濃度50ng/ μ lとなるようにTE緩衝液に溶解し、以下の実施例に使用した。

【0 0 7 5】

＜4. DNAマイクロアレイの作製＞

＜4-1. ガラス基板の洗浄＞

合成石英のガラス基板 (サイズ: 25mm x 75mm x 1mm、飯山特殊ガラス社製) を耐熱、耐アルカリ のラックに入れ、所定の濃度に調製した超音波洗浄用の洗浄液に浸した。一晩洗浄液中で浸した後、20分間超音波洗浄を行った。続いて基板を取り出し、軽く純水ですすいだ後、超純水中で20分超音波洗浄をおこなった。次に80℃に加熱した1N水酸化ナトリウム水溶液中に10分間基板を浸した。再び純水洗浄と超純水洗浄を行い、DNAチップ用の石英ガラス基板を用意した。

【0 0 7 6】

〈 4 - 2 . 表面処理 〉

シランカップリング剤KBM-603(信越シリコン社製)を、1%の濃度となるように純水中に溶解させ、2時間室温で攪拌した。続いて、先に洗浄したガラス基板をシランカップリング剤水溶液に浸し、20分間室温で放置した。ガラス基板を引き上げ、軽く純水で表面を洗浄した後、窒素ガスを基板の両面に吹き付けて乾燥させた。次に乾燥した基板を120℃に加熱したオーブン中で1時間ベークし、カップリング剤処理を完結させ、基板表面にアミノ基を導入した。次いで同仁化学研究所社製のN-マレイミドカプロイロキシスクシイミド(N-(6-Maleimidocaproyloxy)succinimido) (以下EMCSと略す) を、ジメチルスルホキシドとエタノールの1:1混合溶媒中に最終濃度が0.3mg/mlとなるように溶解したEMCS溶液を用意した。ベークの終了したガラス基板を放冷し、調製したEMCS溶液中に室温で2時間浸した。この処理により、シランカップリング剤によって表面に導入されたアミノ基とEMCSのスクシイミド基が反応し、ガラス基板表面にマレイミド基が導入された。EMCS溶液から引き上げたガラス基板を、先述のMCSを溶解した混合溶媒を用いて洗浄し、さらにエタノールにより洗浄した後、窒素ガス雰囲気下で乾燥させた。

【 0 0 7 7 】

〈 4 - 3 . プローブDNA 〉

実施例1で作製した微生物検出用プローブを純水に溶解し、それぞれ、最終濃度(インク溶解時)10 μ Mとなるように分注した後、凍結乾燥を行い、水分を除いた。

【 0 0 7 8 】

〈 4 - 4 . BJプリンターによるDNA吐出、および基板への結合 〉

グリセリン7.5wt%、チオジグリコール7.5wt%、尿素7.5wt%、アセチレノールEH(川研ファインケミカル社製)1.0wt%を含む水溶液を用意した。続いて、先に用意した7種類のプローブ(表1)を上記の混合溶媒に規定濃度なるように溶解した。得られたDNA溶液をバブルジェット(登録商標)プリンター(商品名:BJF-850 キヤノン社製)用インクタンクに充填し、印字ヘッドに装着した。

【 0 0 7 9 】

なおここで用いたバブルジェット（登録商標）プリンターは平板への印刷が可能のように改造を施したものである。またこのバブルジェット（登録商標）プリンターは、所定のファイル作成方法に従って印字パターンを入力することにより、約5ピコリットルのDNA溶液を約120マイクロメートルピッチでスポッティングすることが可能となっている。

【0080】

続いて、この改造バブルジェット（登録商標）プリンターを用いて、1枚のガラス基板に対して、印字操作を行い、アレイを作製した。印字が確実に行われていることを確認した後、30分間加湿チャンバー内に静置し、ガラス基板表面のマレイミド基と核酸プローブ末端のチオール基とを反応させた。

【0081】

<4-5. 洗浄>

30分間の反応後、100mMのNaClを含む10mMのリン酸緩衝液(pH7.0)により表面に残ったDNA溶液を洗い流し、ガラス基板表面に一本鎖DNAが固定したDNAマイクロアレイを得た。

【0082】

<5. 検体の増幅と標識化（PCR増幅&蛍光標識の取り込み）>

検体となる微生物DNAの増幅、および、標識化反応を以下に示す。

【0083】

Premix PCR 試薬 (TAKARA ExTaq)	25	μl
Template Genome DNA	2	μl (100ng)
Forward Primer mix	2	μl (20pmol/tube each)
Reverse Primer mix	2	μl (20pmol/tube each)
Cy-3 dUTP (1mM)	2	μl (2nmol/tube)
H ₂ O	17	μl
Total	50	μl

【0084】

上記組成の反応液を以下のプロトコールに従って、市販のサーマルサイクラーで増幅反応を行った。

【0085】

95°C	10 min.	
92°C	45 sec.	← 35 Cycles →
55°C	45 sec.	
72°C	45 sec.	
72°C	10 min.	

【0086】

反応終了後、精製用カラム (QIAGEN QIAquick PCR Purification Kit) を用いてPrimerを除去した後、増幅産物の定量を行い、標識化検体とした。

【0087】

<6. ハイブリダイゼーション>

<4. DNAマイクロアレイの作製>で作製したDNAマイクロアレイと<5. 検体の増幅と標識化 (PCR増幅&蛍光標識の取り込み)>で作製した標識化検体を用いて検出反応を行った。

【0088】

<6-1. DNAマイクロアレイのブロッキング>

BSA (牛血清アルブミンFraction V: Sigma社製) を1wt%となるように100mM NaCl / 10mM Phosphate Bufferに溶解し、この溶液に<DNAマイクロアレイの作製>で作製したDNAマイクロアレイを室温で2時間浸し、ブロッキングを行った。ブロッキング終了後、0.1wt% SDS (ドデシル硫酸ナトリウム) を含む 2 x SSC溶液 (NaCl 300mM、Sodium Citrate (trisodium citrate dihydrate, $C_6H_5Na_3 \cdot 2H_2O$) 30mM、p.H. 7.0) で洗浄を行った後、純水でリンスしてからスピンドライ装置で水切りを行った。

【 0 0 8 9 】

＜ 6 - 2 . ハイブリダイゼーション ＞

水切りしたDNAマイクロアレイをハイブリダイゼーション装置 (Genomic Solutions Inc. Hybridization Station) にセットし、以下に示すハイブリダイゼーション溶液、条件でハイブリダイゼーション反応を行った。

【 0 0 9 0 】

・ ハイブリダイゼーション溶液

6 x SSPE / 10% Form amide / Target (2nd PCR Products 全量)
(6xSSPE: NaCl 900mM、NaH₂PO₄·H₂O 60mM、EDTA 6mM、p.H. 7.4)

・ ハイブリダイゼーション条件

65 °C 3min → 92°C 2min → 45°C 3hr → Wash 2 x SSC / 0.1% SDS at 25°C → Wash 2 x SSC at 20°C → (Rinse with H₂O : Manual)
→ Spin dry (65°Cで3分、92度で2分、45°Cで3時間ハイブリダイゼーション反応させた後、2 x SSC / 0.1% SDS、25°Cで洗浄、2xSSC、20°Cで洗浄後、純水でリンスしスピンドライした)。

【 0 0 9 1 】

＜ 7 . 微生物の検出 (蛍光測定) ＞

ハイブリダイゼーション反応終了後のDNAマイクロアレイをDNAマイクロアレイ用蛍光検出装置 (Axon社製、GenePix 4000B) を用いて蛍光測定を行った。夫々のプローブで良好な判別結果が得られた。

【 0 0 9 2 】

以上説明したように、本実施形態によれば、DNAマイクロアレイシステムに最適なオリゴヌクレオチドプローブの設計が実現できる。このことにより、より確かな生物種、個体識別情報が得られるという効果がある。

【 0 0 9 3 】

【他の実施形態】

なお、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ (またはCPUやMPU) が記憶媒体に格納

されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。

【0094】

この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【0095】

プログラムコードを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、CD-ROM、CD-R、磁気テープ、不揮発性のメモリカード、ROMなどを用いることができる。

【0096】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているOS（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0097】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるCPUなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【発明の効果】

以上説明したように、本発明によれば、正確かつ、再現性のあるプローブ設計が可能となる。

【配列表】

〈110〉 キヤノン株式会社 CANON, INC.

<120> DNAプローブ設計装置及びDNAプローブ設計のための情報処理方法

<160> 162

<210> 1

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 1

gaaccgcatg gttcaaaagt gaaaga

<210> 2

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 2

cacttataga tggatccgcg ctgc

<210> 3

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 3

tgcacatctt gacggtacct aatcag

<210> 4

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 4

ccccttagtg ctgcagctaa cg

<210> 5

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 5

aatacaaagg gcagcgaaac cgc

<210> 6

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 6

ccggtggagt aaccttttag gagct

<210> 7

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 7

taacctttta ggagctagcc gtcga

<210> 8

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 8

tttaggagct agccgtcgaa ggt

<210> 9

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 9

tagccgtcga aggtgggaca aat

<210> 10

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 10

gaacagacga ggagcttgct cc

<210> 11

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 11

tagtgaaaga cggttttgct gtcact

<210> 12

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 12

taagtaacta tgcacgtctt gacggt

<210> 13

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 13

gacccctcta gagatagagt tttccc

<210> 14

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 14

agtaaccatt tggagctagc cgtc

<210> 15

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 15

gagcttgctc ctctgacgtt agc

<210> 16

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 16

agccggtgga gtaaccattt gg

<210> 17

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 17

ctcttgccat cggatgtgcc ca

<210> 18

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 18

atacctttgc tcattgacgt tacccg

<210> 19

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 19

tttgctcatt gacgttacc gcag

<210> 20

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 20

actggcaagc ttgagtctcg taga

<210> 21

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 21

atacaaagag aagcgacctc gcg

<210> 22

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 22

cggacctcat aaagtgcgtc gtagt

<210> 23

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 23

gcgggggagga agggagtaaa gttaat

<210> 24

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 24

tagcacagag agcttgctct cgg

<210> 25

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 25

tcatgccatc agatgtgccc aga

<210> 26

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 26

cggggaggaa ggcgataagg ttaat

<210> 27

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 27

ttcgattgac gttacccgca gaaga

<210> 28

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 28

ggtctgtcaa gtcggatgtg aaatcc

<210> 29

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 29

gcaggctaga gtctttaga gggg

<210> 30

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 30

tgagggagaa agtgggggat cttc

<210> 31

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 31

tcagatgagc ctaggtcgga ttagc

<210> 32

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 32

gagctagagt acggtagagg gtgg

<210> 33

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 33

gtacggtaga gggtggtgga atttc

<210> 34

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 34

gaccacctgg actgatactg acac

<210> 35

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 35

tggccttgac atgctgagaa ctttc

<210> 36

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 36

ttagttacca gcacctcggg tgg

<210> 37

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 37

tagtctaacc gcaaggggga cg

<210> 38

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 38

tagcacaggg agcttgctcc ct

<210> 39

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 39

aggtggtgag cttaatacgc tcac

<210> 40

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 40

tcaccaattg acgttactcg cagaag

<210> 41

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 41

actgcatttg aaactggcaa gctaga

<210> 42

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 42

ttatcctttg ttgcagcttc ggcc

<210> 43

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 43

actttcagcg aggaggaagg tgg

<210> 44

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 44

agtagaacgc tgaaggagga gcttg

<210> 45

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 45

cttgcacac taccagatgg acctg

<210> 46

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 46

tgagagtgga aagttcacac tgtgac

<210> 47

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 47

gctgtggctt aaccatagta ggcttt

<210> 48

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 48

aagcggctct ctggcttgta act

<210> 49

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 49

tagacccttt ccggggttta gtgc

<210> 50

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 50

gacggcaagc taatctctta aagcca

<210> 51

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 51

gcttgggaat ctggcttatg gagg

<210> 52

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 52

tgccatagga tgagcccaag tgg

<210> 53

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 53

cttgggaatg tactgacgct catgtg

<210> 54

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 54

ggattgggct tagagcttgg tgc

<210> 55

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 55

tacagaggga agcgaagctg cg

<210> 56

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 56

ggcgtttacc acggtatgat tcatga

<210> 57

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 57

aatgcctacc aagcctgcga tct

<210> 58

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 58

tatcggaaga tgaaagtgcg ggact

<210> 59

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 59

cagagagctt gctctcgggt ga

<210> 60

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 60

gggaggaagg tgttgtggtt aataac

<210> 61

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 61

ggtgttgtgg ttaataacca cagcaa

<210> 62

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 62

gcggtctgtc aagtcggatg tg

<210> 63

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 63

attcgaaact ggcaggctag agtct

<210> 64

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 64

taaccacagc aattgacgtt acccg

<210> 65

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 65

gcaattgacg ttacccgcag aaga

<210> 66

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 66

ttctttcctc ccgagtgtt gca

<210> 67

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 67

aacacgtggg taacctaccc atcag

<210> 68

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 68

atggcataag agtgaaaggc gctt

<210> 69

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 69

gacccgcggt gcattagcta gt

<210> 70

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 70

ggacgttagt aactgaacgt cccct

<210> 71

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 71

ctcaaccggg gagggtcatt gg

<210> 72

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 72

ttggagggtt tccgcccttc ag

<210> 73

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 73

gcggcgtgcc taatacatgc aag

<210> 74

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 74

gcggcaggcc taacacatgc aag

<210> 75

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 75

gcggcaggct taacacatgc aag

<210> 76

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 76

atccagccgc accttccgat ac

<210> 77

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 77

atccaaccgc aggttcccct ac

<210> 78

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 78

atccagccgc aggttcccct ac

<210> 79

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 79

tctttcactt ttgaaccatg cggttc

<210> 80

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 80

gcagcgcgga tccatctata agtg

<210> 81

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 81

ctgattaggt accgtcaaga tgtgca

<210> 82

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 82

cgttagctgc agcactaagg gg

<210> 83

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 83

gcggtttcgc tgccctttgt att

<210> 84

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 84

agctcctaaa aggttactcc accgg

<210> 85

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 85

tcgacggcta gctcctaaaa ggtta

<210> 86

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 86

accttcgacg gctagctcct aaa

<210> 87

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 87

atttgtccca ccttcgacgg cta

<210> 88

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 88

ggagcaagct cctcgtctgt tc

<210> 89

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 89

agtgacagca aaaccgtctt tcacta

<210> 90

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 90

accgtcaaga cgtgcatagt tactta

<210> 91

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 91

gggaaaactc tatctctaga ggggtc

<210> 92

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 92

gacggctagc tccaaatggt tact

<210> 93

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 93

gctaacgtca gaggagcaag ctc

<210> 94

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 94

ccaaatgggtt actccaccgg ct

<210> 95

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 95

tgggcacatc cgatggcaag ag

<210> 96

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 96

cgggtaacgt caatgagcaa aggtat

<210> 97

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 97

ctgcgggtaa cgtcaatgag caaa

<210> 98

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 98

tctacgagac tcaagcttgc cagt

<210> 99

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 99

cgcgaggtcg cttctctttg tat

<210> 100

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 100

actacgacgc actttatgag gtccg

<210> 101

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 101

attaacttta ctcccttcct ccccgc

<210> 102

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 102

ccgagagcaa gctctctgtg cta

<210> 103

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 103

tctgggcaca tctgatggca tga

<210> 104

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 104

attaacctta tcgccttcct ccccg

<210> 105

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 105

tcttctgcgg gtaacgtcaa tcgaa

<210> 106

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 106

ggatttcaca tccgacttga cagacc

<210> 107

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 107

cccctctaca agactctagc ctgc

<210> 108

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 108

gaagatcccc cactttctcc ctca

<210> 109

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 109

gctaatacga cctaggctca tctga

<210> 110

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 110

ccaccctcta ccgtactcta gctc

<210> 111

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 111

gaaattccac caccctctac cgtac

<210> 112

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 112

gtgtcagtat cagtccaggt ggtc

<210> 113

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 113

gaaagttctc agcatgtcaa ggcca

<210> 114

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 114

ccacccgagg tgctggtaac taa

<210> 115

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 115

cgteccccctt gcggttagac ta

<210> 116

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 116

agggagcaag ctccctgtgc ta

<210> 117

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 117

gatgagcgta ttaagctcac cacct

<210> 118

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 118

cttctgcgag taacgtcaat tgatga

<210> 119

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 119

tctagcttgc cagtttcaaa tgcagt

<210> 120

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 120

ggccgaagct gcaacaaagg ataa

<210> 121

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 121

ccaccttcct cctcgctgaa agt

<210> 122

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 122

caagctcctc cttcagcggt ctact

<210> 123

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 123

caggtccatc tggtagtgat gcaag

<210> 124

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 124

gtcacagtgt gaactttcca ctctca

<210> 125

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 125

aaagcctact atggttaagc cacagc

<210> 126

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 126

agttacaagc cagagagccg ctt

<210> 127

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 127

atctgggaaa ggccccaaat cacg

<210> 128

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 128

tggtttaag agattagctt gccgtc

<210> 129

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 129

cctccataag ccagattccc aagc

<210> 130

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 130

ccacttgggc tcacccatg gca

<210> 131

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 131

cacatgagcg tcagtacatt cccaag

<210> 132

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 132

gcaccaagct ctaagcccaa tcc

<210> 133

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 133

cgcagcttcg cttccctctg ta

<210> 134

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 134

tcatgaatca taccgtggta aacgcc

<210> 135

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 135

agatcgcagg cttggtaggc att

<210> 136

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 136

agtcccgcac tttcatcttc cgata

<210> 137

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 137

tgcaagcact cgggaggaaa gaa

<210> 138

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 138

ctgatgggta ggttaccac gtgtt

<210> 139

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 139

aagcgcccttt cactcttatg ccat

<210> 140

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 140

actagctaat gcaccgcggg tc

<210> 141

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 141

aggggacgtt cagttactaa cgtcc

<210> 142

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 142

ccaatgaccc tccccggttg ag

<210> 143

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 143

ctgaagggcg gaaaccctcc aa

<210> 144

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 144

tcacccgaga gcaagctctc tg

<210> 145

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 145

gttattaacc acaacacctt cctccc

<210> 146

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 146

ttgctgtggt tattaaccac aacacc

<210> 147

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 147

cacatccgac ttgacagacc gc

<210> 148

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 148

agactctagc ctgccagttt cgaat

<210> 149

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 149

cgggtaacgt caattgctgt ggтта

<210> 150

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 150

tctttctgcgg gtaacgtcaa ttgc

<210> 151

<211> 10

<212> DNA

<213> Artificial Sequence

<400> 151

tgatgtgaaa

<210> 152

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 152

gcccacggct caaccgtgga ggg

<210> 153

<211> 10

<212> DNA

<213> Artificial Sequence

<400> 153

aaactggaaa

<210> 154

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 154

gaaccgcatg gttcaaaagt gaaaga

<210> 155

<211> 24

<212> DNA

<213> Artificial Sequence

<400> 155

cacttataga tggatccgcg ctgc

<210> 156

<211> 26

<212> DNA

<213> Artificial Sequence

<400> 156

tgcacatctt gacgtacct aatcag

<210> 157

<211> 22

<212> DNA

<213> Artificial Sequence

<400> 157

ccccttagtg ctgcagctaa cg

<210> 158

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 158

aatacaaagg gcagcgaaac cgc

<210> 159

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 159

ccggtggagt aaccttttag gagct

<210> 160

<211> 25

<212> DNA

<213> Artificial Sequence

<400> 160

taacctttta ggagctagcc gtcga

<210> 161

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 161

tttaggagct agccgtcgaa ggt

<210> 162

<211> 23

<212> DNA

<213> Artificial Sequence

<400> 162

tagccgtcga aggtgggaca aat

【図面の簡単な説明】

【図 1】

実施形態によるプローブ設計方法の概要を説明する図である。

【図 2】

実施形態によるプローブ設計方法が適用可能な情報処理装置の構成を示すブロック図である。

【図 3】

ハイブリダイゼーション反応を説明する図である。

【図 4】

感染症の判定用 DNA マイクロアレイの実験手順を説明する図である。

【図 5】

黄色ブドウ球菌の異なる株のゲノム構造を説明する図である。

【図 6】

実施形態による頻度テーブルの例を示す図である。

【図 7】

ターゲット塩基配列の特異性をスキャンした例を示す図である。

【図 8】

競合頻度テーブルと自己頻度テーブルの値をグラフ化した図である。

【図 9】

実施形態によるプローブ設計方法を説明するフローチャートである。

【図 1 0】

実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図 1 1】

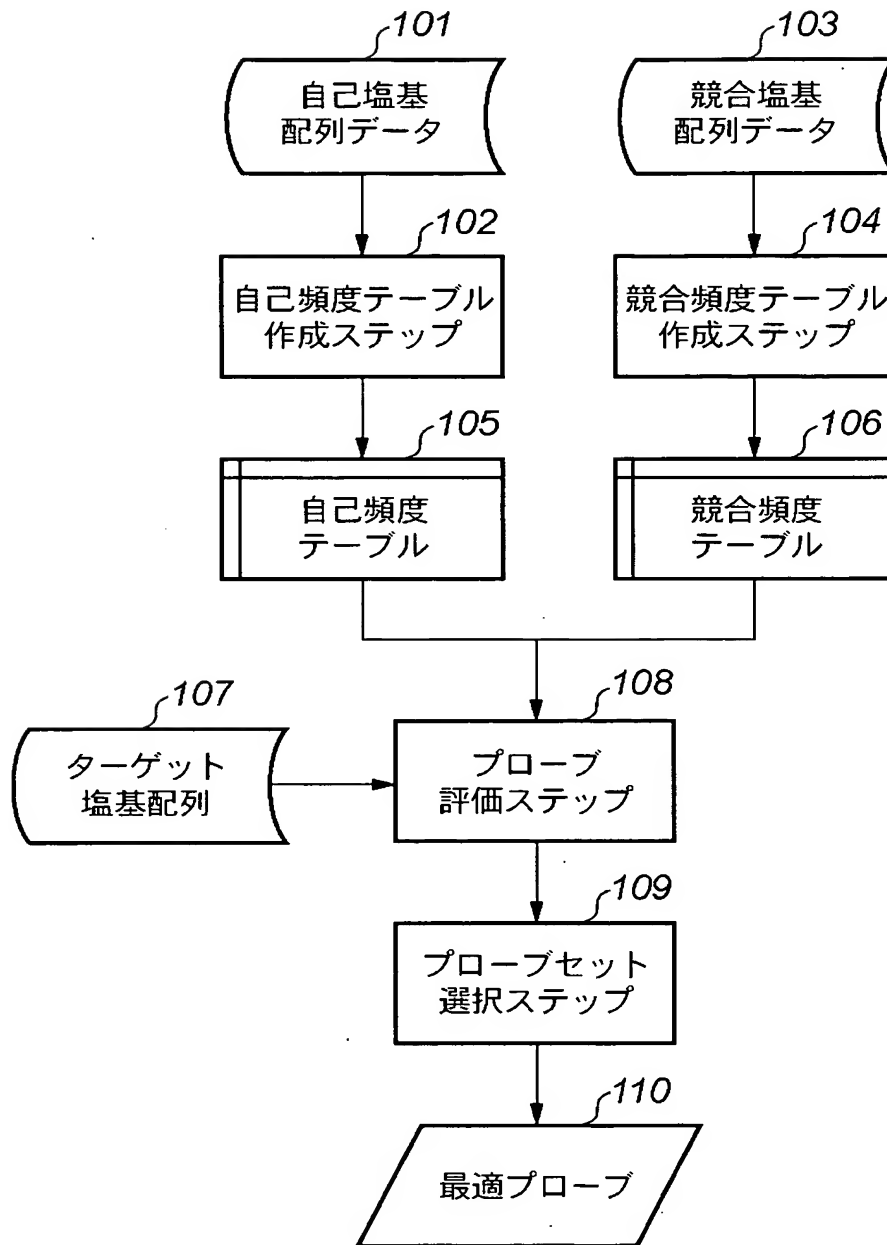
実施形態によるプローブ設計方法におけるユーザインターフェースを示す図である。

【図 1 2】

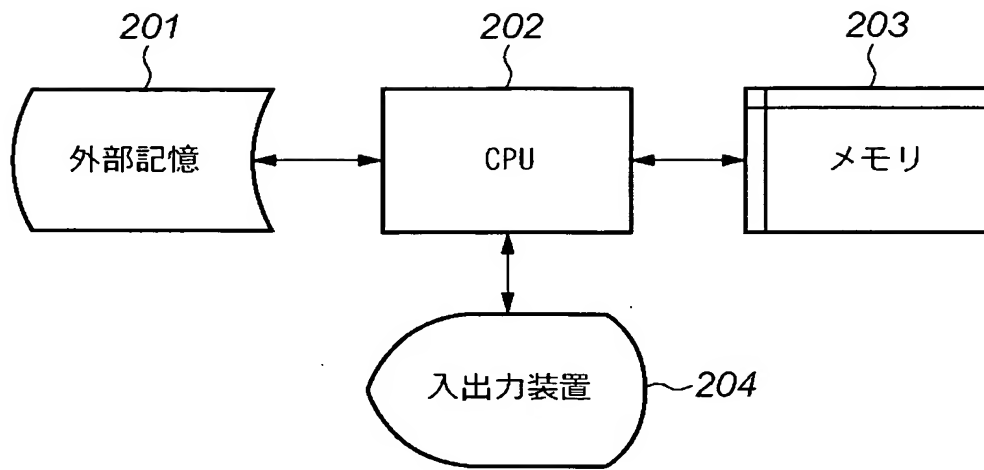
実施形態のプローブ設計方法で設計したプローブを用いて核酸解析を行う手順を説明する図である。

【書類名】 図面

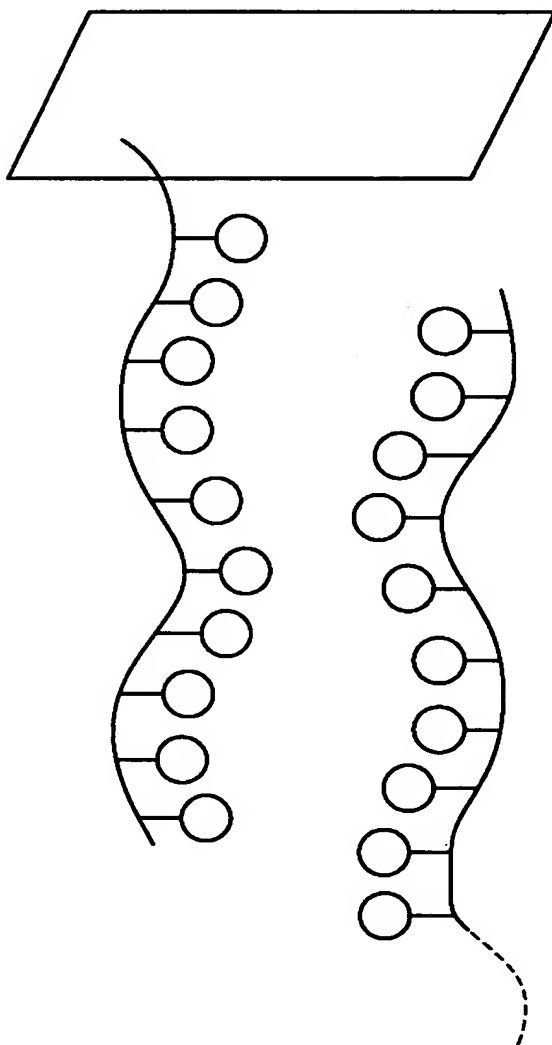
【図 1】



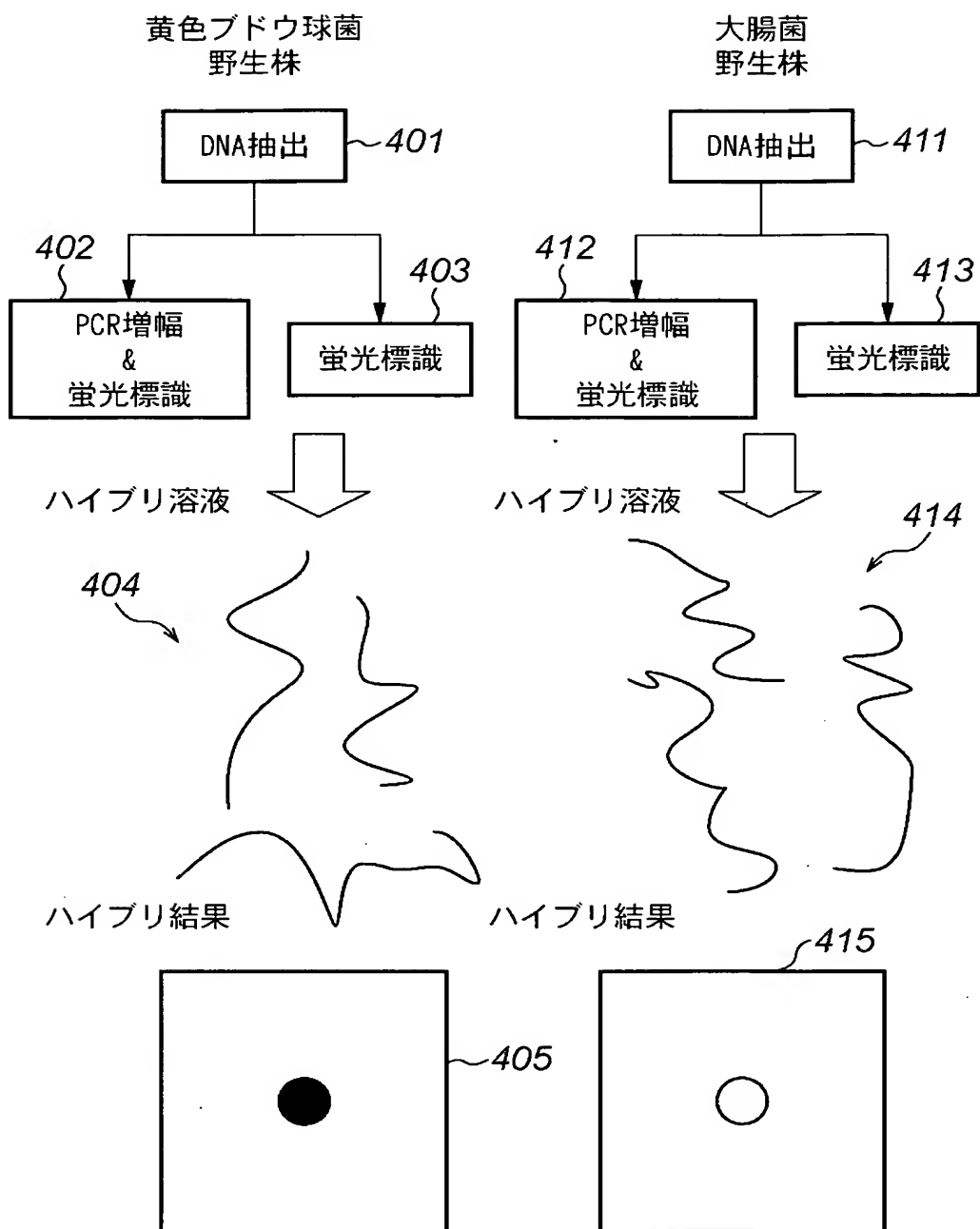
【図 2】



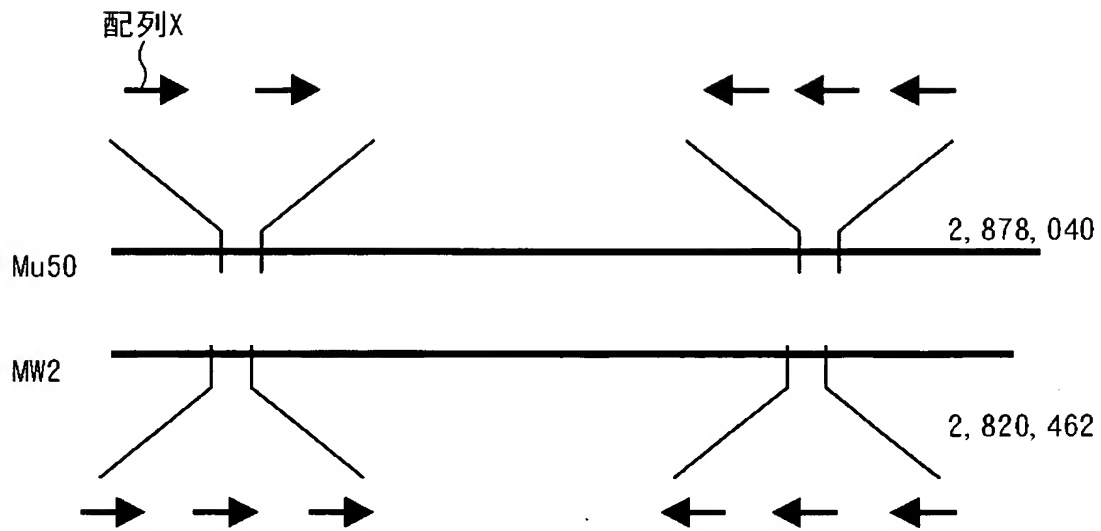
【図 3】



【図 4】



【図 5】



【図 6】

塩基配列	頻度(一特異性)
AAAAAAAAA	1000
AAAAAAAAAT	10
...	...
...	...
GTAGATCCA	3
GTAGATCCT	5
...	...
...	...
CCCCCCCCC	0

【図 7】

ターゲット塩基配列

ATCGTA.....GTAA

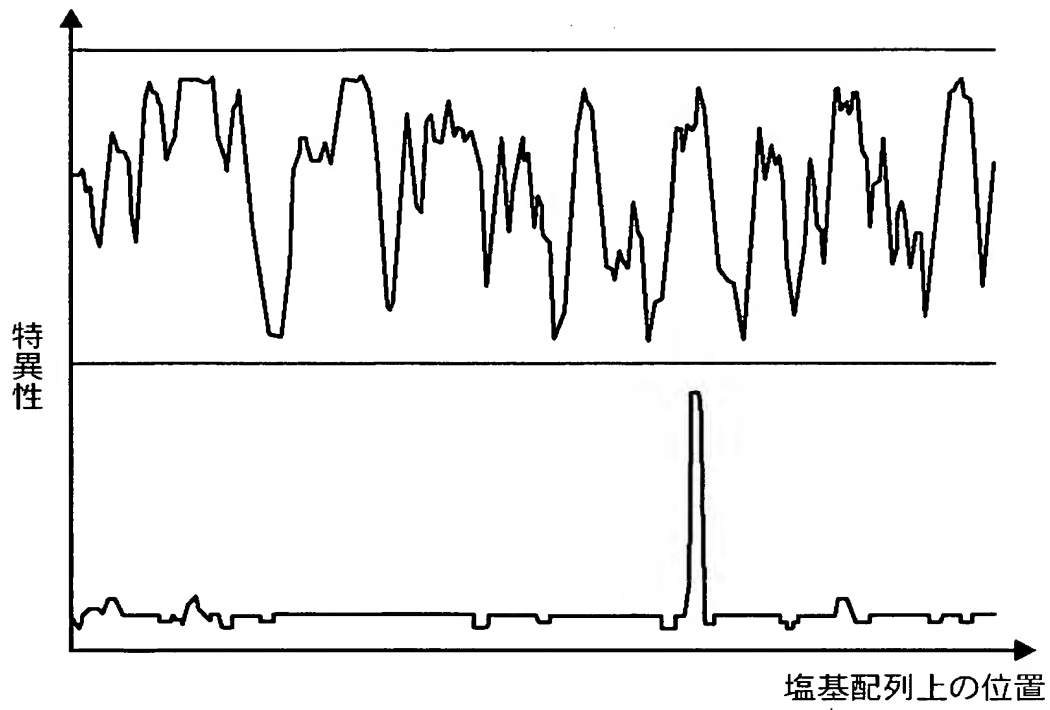
特異性

ATCGTACG
TGGTACGA
GCGTACGAC
CGTACGACT

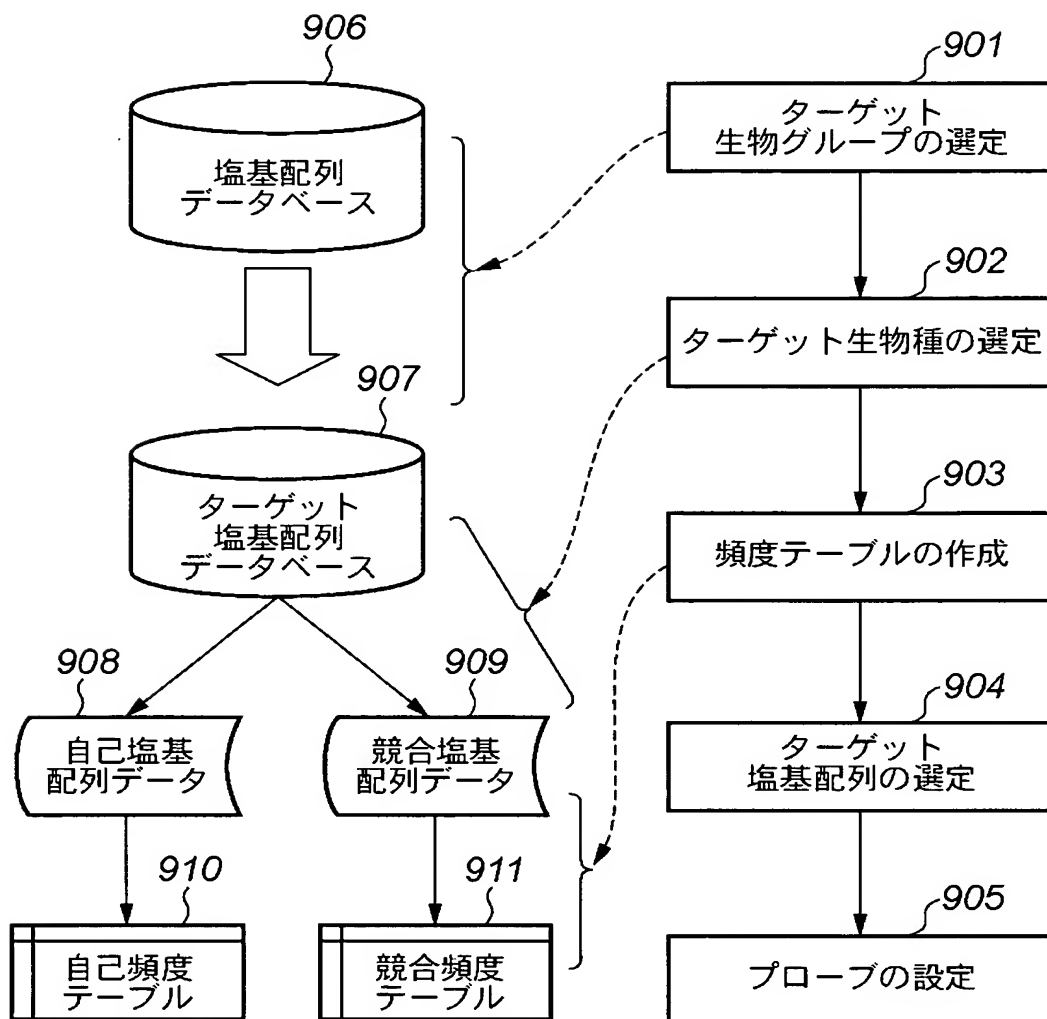
-300
-250
-200
-220
-400
...
...
-300
-350

TATGCGTAA
ATGCGTAAA

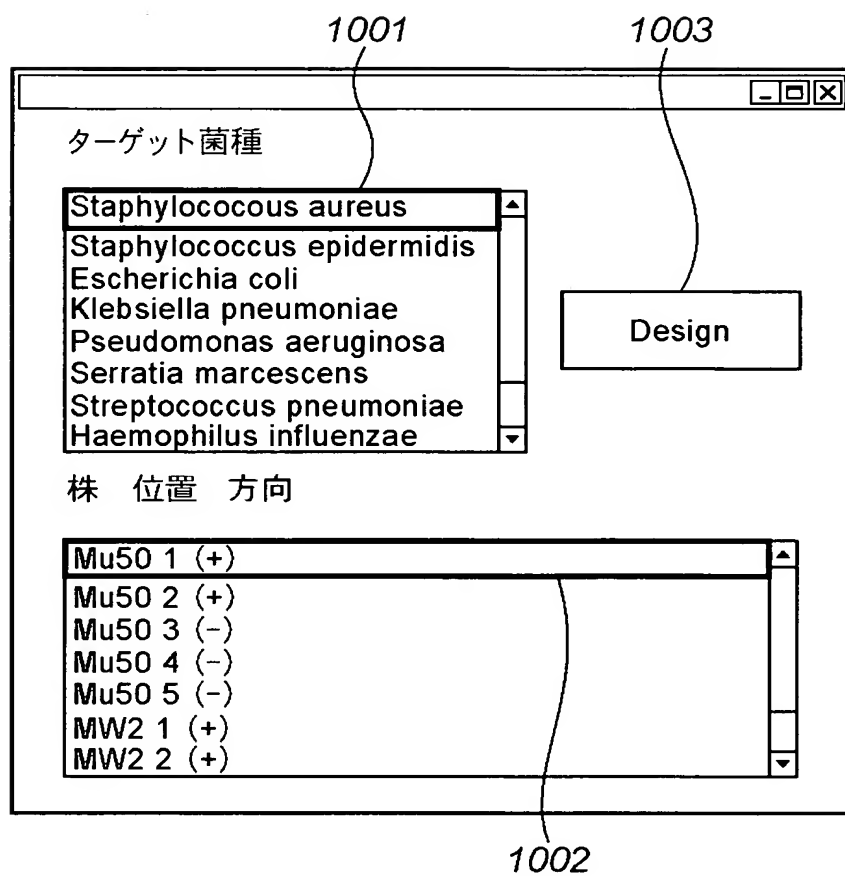
【図 8】



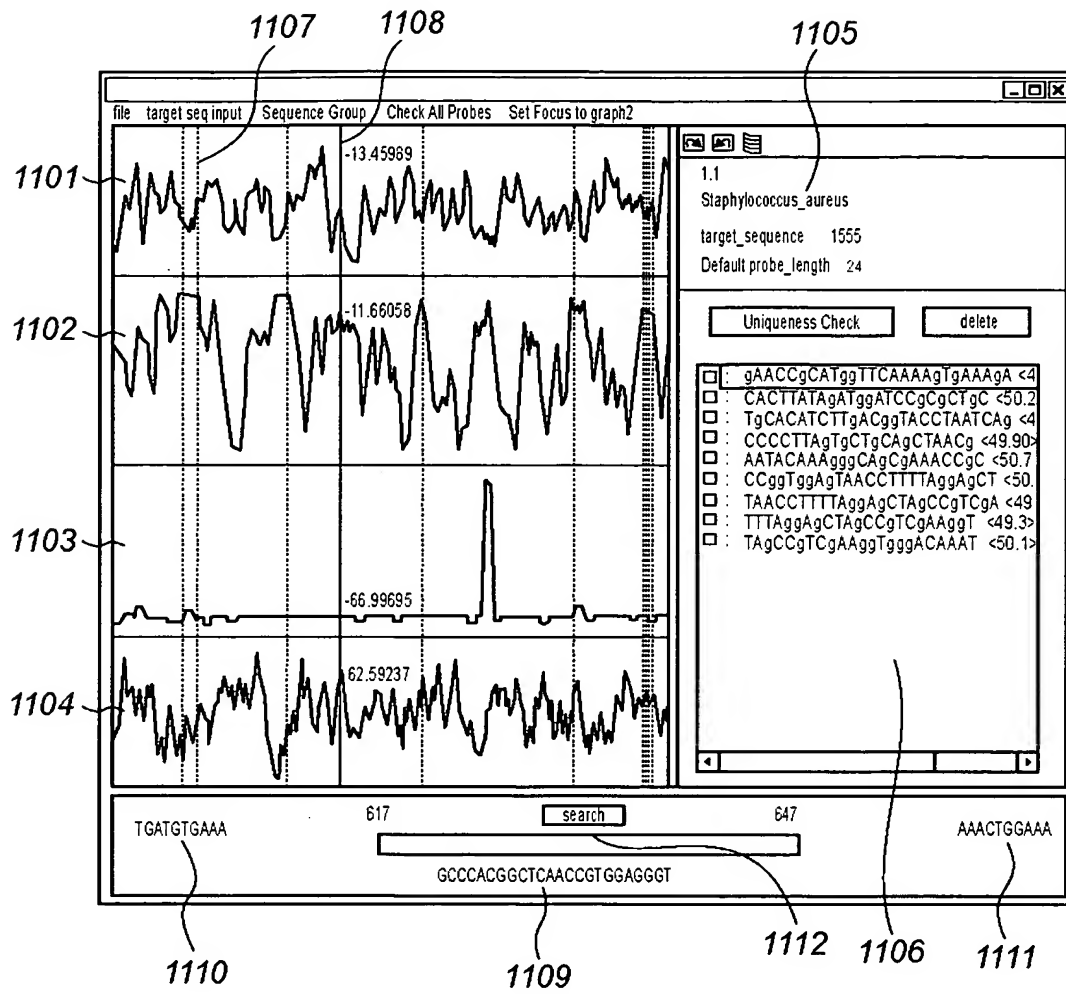
【図 9】



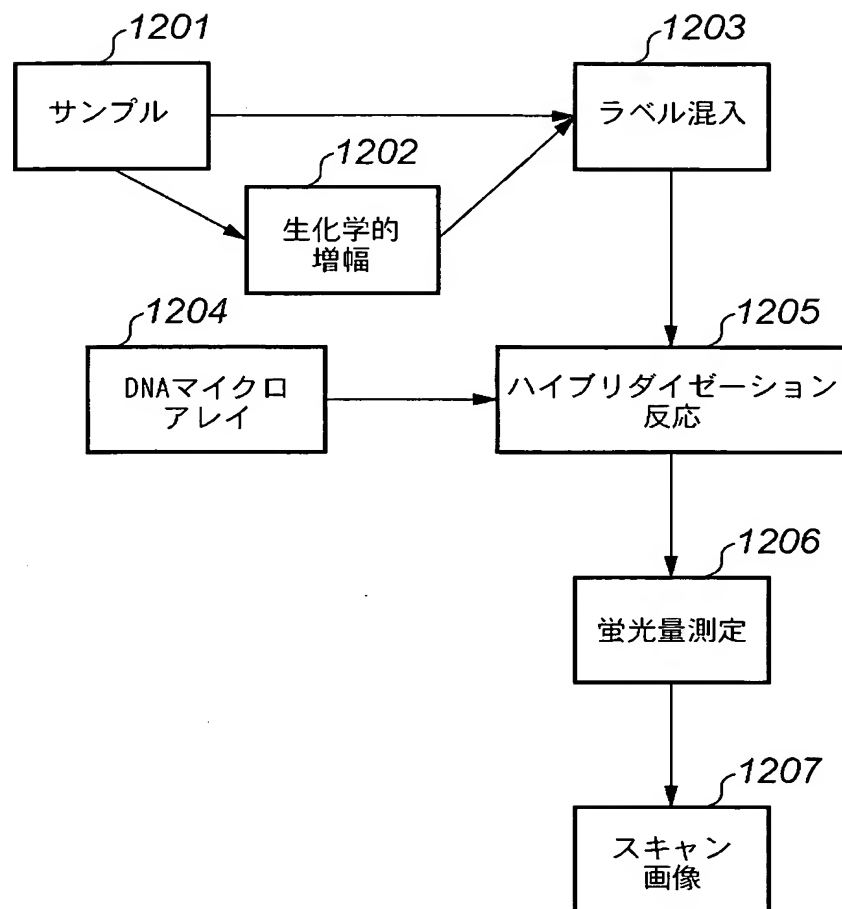
【図 10】



【図 11】



【図 12】



【書類名】 要約書

【要約】

【課題】 正確かつ、再現性のあるDNAプローブ設計を可能とする。

【解決手段】 ターゲット塩基配列を含む自己塩基配列データ101について、ターゲット塩基配列107から得られる複数の部分塩基配列の各々の出現数をカウントして得られる自己頻度テーブル105を保持する。また、自己塩基配列データ101とは区別されるべき競合塩基配列データ103について、上記複数の部分塩基配列の各々の出現数をカウントして得られる競合頻度テーブルを保持する。プローブ評価ステップ108では、自己及び競合頻度テーブル105、106の頻度情報を部分塩基配列を基準に対比可能に提示し、ユーザによる指示操作に応じて複数の部分塩基配列の少なくとも一つを特定し、特定された部分塩基配列に基づいてプローブ候補が形成される。

【選択図】 図1

特願 2 0 0 3 - 0 9 9 4 6 4

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 1 0 0 7]

1. 変更年月日 1 9 9 0 年 8 月 3 0 日

[変更理由] 新規登録

住 所 東京都大田区下丸子 3 丁目 3 0 番 2 号

氏 名 キヤノン株式会社